

Monte Carlo generators for top quark physics at the LHC

B.S. ACHARYA⁽¹⁾, F. CAVALLARI⁽²⁾, G. CORCELLA^{(3)(4)(*)}, R. DI SIPIO⁽⁵⁾ and G. PETRUCCIANI⁽⁴⁾

⁽¹⁾ *Abdul Salam International Center for Theoretical Physics and INFN, Sezione di Trieste, Italy*

⁽²⁾ *INFN, Sezione di Roma, Italy*

⁽³⁾ *Museo Storico della Fisica e Centro Studi e Ricerche E. Fermi, Roma, Italy*

⁽⁴⁾ *Scuola Normale Superiore and INFN, Sezione di Pisa, Italy*

⁽⁵⁾ *Università di Bologna and INFN, Sezione di Bologna, Italy*

Summary. — We review the main features of Monte Carlo generators for top quark phenomenology and present some results for $t\bar{t}$ and single-top signals and backgrounds at the LHC.

PACS 14.65.Ha – Top quarks.

PACS 12.38.Bx – Perturbative calculations.

1. – Introduction

In order to perform precise measurements of properties of top quarks at the LHC, the use of reliable tools will be essential. Extensive work has been carried out to improve parton-shower and matrix-element Monte Carlo (MC) codes for $t\bar{t}$ and single-top signals and backgrounds. Parton shower generators, such as HERWIG [1] or PYTHIA [2], simulate top production and decay using leading-order (LO) matrix elements and describe multiple radiation in the soft or collinear approximation. HERWIG showers satisfy angular ordering, valid in the soft approximation after azimuthal averaging, while PYTHIA orders its cascades according to transverse momentum or virtuality, with an option to veto non-angular-ordered emissions. To simulate hard and large-angle radiation, both PYTHIA and HERWIG have been provided with matrix-element corrections. HERWIG splits the physical phase space into a soft/collinear region, where one trusts parton showers, and a region, corresponding to hard and large-angle radiation, where the exact amplitude is used. Moreover, HERWIG corrects the ‘hardest-so-far’ emission in the shower. The PYTHIA implementation is somewhat different: the parton shower

(*) Talk given by G. Corcella at ‘V Workshop Italiano sulla Fisica pp a LHC’, Perugia, Italy, 30 January – 2 February, 2008.

approximation is used throughout the entire physical phase space and the first emission is generated exactly. HERWIG [3, 4] and PYTHIA [5] include matrix-element corrections to top-quark decay, but not to top production. Such corrections are also applied to W -boson production, which is matched to the $W + 1$ jet process [6, 7], one of the backgrounds for top production. At the end of the parton cascade, HERWIG and PYTHIA simulate the hadronization transition according to the cluster [8] and string [9] models, respectively. As for the underlying event, both programs deal with it by allowing multiple interactions [10, 11]. For this purpose, HERWIG is interfaced with the JIMMY code. HERWIG++ [12] and PYTHIA 8 [13] are the new object-oriented versions of the two codes, written in C++. PYTHIA 8 implements only transverse-momentum evolution, with initial/final-state radiation and multiple interactions interweaved in a common ordering. HERWIG++ includes a mass-dependent term in the splitting functions, leading to a better treatment of the radiation off heavy quarks and heavy-hadron fragmentation.

Besides HERWIG and PYTHIA, other parton showers generators are available. The ARIADNE [14] code uses PYTHIA hard scattering and hadronization, but its showers are dipole cascades, i.e. $2 \rightarrow 3$ branchings in the soft limit. The evolution variable of ARIADNE is transverse momentum: in fact, implementing dipole emissions, there is no need for azimuthal averaging and angular ordering. The SHERPA [15] generator implements showers which, like in PYTHIA, are ordered in virtuality, with an option to reject non-angular-ordered emissions. Hadronization occurs following the string model. A feature of ARIADNE and SHERPA is that matrix-element matching is implemented according to the CKKW-L [16, 17] method. Jets are clustered according to the k_T -algorithm [18], with a threshold y_{cut} on the k_T -clustering variable y_{ij} . Exact matrix elements, weighted by the Sudakov form factors, and parton-shower approximations are then used above and below y_{cut} .

The programs mentioned so far, even after matrix-element matching, still yield LO rates. Next-to-leading order (NLO) cross sections and observables are instead given by the codes MC@NLO [19] and POWHEG [20], both available for top production. MC@NLO implements the hard scattering at NLO and employs HERWIG for showers and hadronization: its drawbacks are its shower-model (HERWIG) dependence and its simulation of events with negative weights. Such problems have been overcome by POWHEG, which can be interfaced to any shower model and does not have any negatively-weighted event. The first POWHEG emission is the hardest in transverse momentum (p_T) and is generated exactly at NLO; the subsequent cascade is ordered in p_T . To include angular ordering, one would need truncated showers, whose implementation is in progress.

Finally, in order to simulate multi-jet final states in top-quark events and backgrounds, such as $W/Z +$ jets, the best available tools are the so-called ‘matrix-element generators’, such as ALPGEN [21], MadGraph/MadEvent [22], CompHep [23], HELAC [24], GR@PPA [25] and WHIZARD [26], which include LO multi-parton amplitudes. Such programs can be interfaced to HERWIG or PYTHIA for showering and hadronization, according to the Les Houches accord. The ALPGEN code, for example, includes the following processes ($V = W$ or Z): $t\bar{t} + 6$ jets, $V + 6$ jets, $Wb\bar{b} + 6$ jets, $VV + 3$ jets. It uses the so-called MLM prescription to match its partons with the subsequent showers. Roughly speaking, one starts, e.g., with $W + n$ partons and defines a cone jet algorithm. After showering, a k -jet event is accepted only if $k \geq n$ and each of the n original partons is clustered into a different jet.

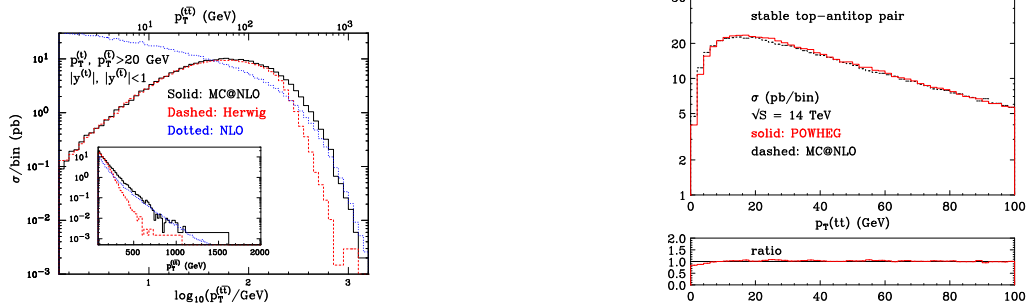


Fig. 1.: Transverse momentum of the $t\bar{t}$ pair at the LHC, according to HERWIG, MC@NLO and the NLO calculation (left), and according to POWHEG and MC@NLO, for stable top quarks (right).

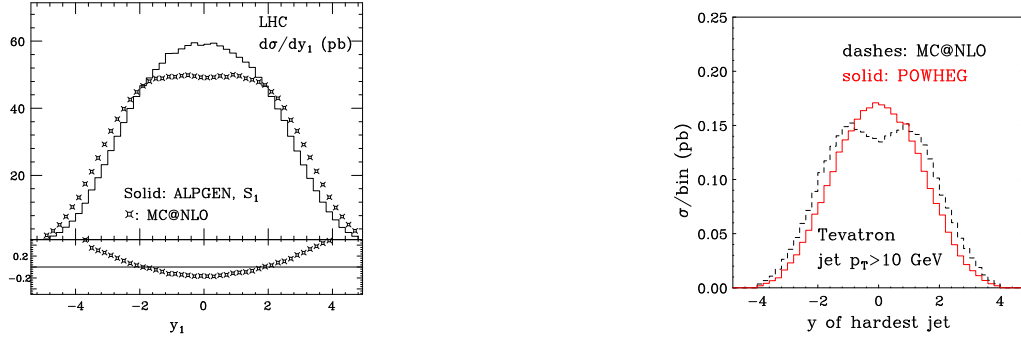


Fig. 2.: Rapidity of the hardest jet in $t\bar{t}(g)$ events at the LHC according to ALPGEN and MC@NLO (left), and at the Tevatron according to POWHEG and MC@NLO (right).

2. – Results for top-quark signals and backgrounds

We would like to present some results for top-quark signals and backgrounds at the LHC, using the programs described above. Following [27], in Fig. 1 we present the transverse momentum of the $t\bar{t}$ pair at the LHC according to HERWIG, MC@NLO and the parton-level NLO calculation, with all distributions normalized to the same area. At small $p_T^{(t\bar{t})}$, where one is mostly sensitive to soft/collinear parton radiation, MC@NLO and HERWIG agree, while the pure NLO calculation is far above the two codes. At large $p_T^{(t\bar{t})}$, it is instead the NLO computation which is reliable: MC@NLO and the NLO prediction agree, while HERWIG underestimates the large- $p_T^{(t\bar{t})}$ rate. In Fig. 1 we also compare MC@NLO and POWHEG for stable top-quark production: up to a small discrepancy at very low $p_T^{(t\bar{t})}$, the agreement between the two codes is remarkable [28]. In Fig. 2 we present the rapidity of the hardest jet y_1 , yielded by MC@NLO and ALPGEN: we have clear disagreement around $y_1 = 0$, with ALPGEN giving more events than MC@NLO. As pointed out in [29], this is due to the fact that MC@NLO does not generate enough hard scatterings with $t\bar{t}g$ in the final state. This is confirmed by the fact that POWHEG, whose first emission is always generated at NLO, yields a y_1 -spectrum similar to ALPGEN (see Fig. 2).

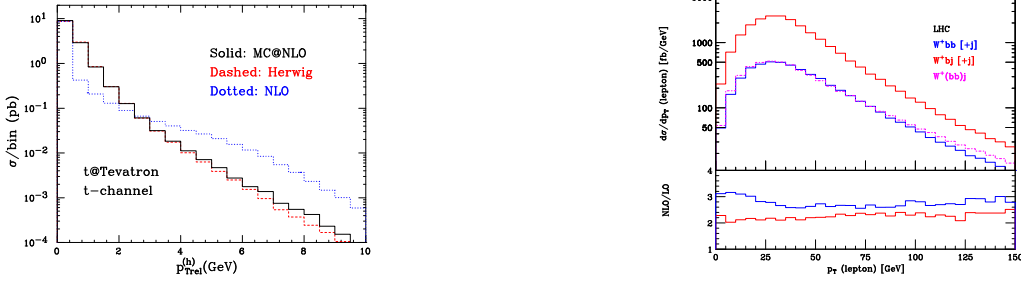


Fig. 3.: Left: relative transverse momentum of the hardest-jet partons with respect to the jet axis for single-top production at the Tevatron. Right: lepton p_T for $W + 2$ jets with a b tag, with the W decaying leptonically, at the LHC. The highest cross section is yielded by the $Wbj[+j]$ subprocesses (solid), with $Wbb[+j]$ (solid) and $W(bb)j$ (dashes) giving similar contributions. The NLO/LO ratio is higher for $Wbb[+j]$.

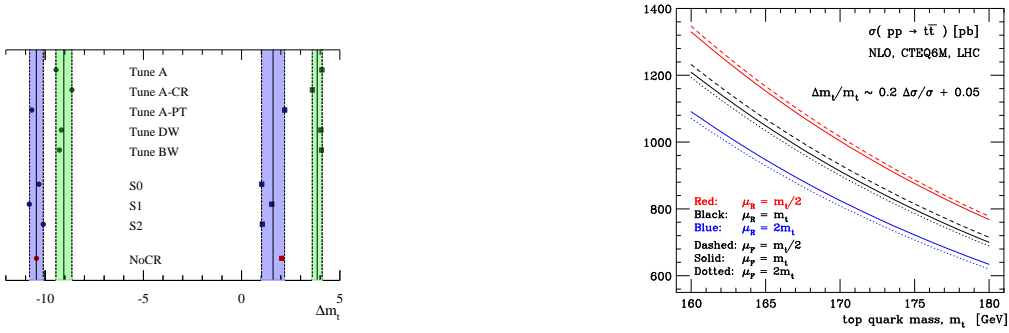


Fig. 4.: Left: uncertainty Δm_t due to the underlying event in PYTHIA. The bands to the $\Delta m_t > 0$ side account for jet-energy-scale corrections. In both cases, results on the left- (right-) hand side correspond to evolution in transverse momentum (virtuality). Right: Dependence of the $t\bar{t}$ cross section on m_t for few values of renormalization and factorization scales. The lines on the top correspond to $\mu_R = m_t/2$, in the middle to $\mu_R = m_t$, on the bottom to $\mu_R = 2m_t$.

For single-top production, a comparison among some of the available codes was performed in [30]. Fig. 3 shows the total transverse momentum of all partons in the hardest jet, relative to the jet axis, using HERWIG, MC@NLO and the NLO calculation. The agreement between HERWIG and MC@NLO is acceptable over the full range: in fact, $p_{T\text{rel}}^h$ is a sufficiently inclusive observable, so that NLO effects mainly result in an overall K -factor. The NLO result is sharply peaked at $p_{T\text{rel}}^h = 0$, since at NLO a jet often coincides with a single parton. At large $p_{T\text{rel}}^h$, the NLO result overestimates the prediction of the two MC codes, where a large- $p_{T\text{rel}}^h$ hadron likely belongs to another jet. Among the backgrounds to single-top production, $W + 2$ jets, at least one with a b quark, were studied in [32], making use of MCFM [31], a NLO parton-level code. Fig. 3 shows the p_T distribution of the lepton from W decay, for the $Wb\bar{b} + X$, $Wbj + X$ and $W(bb)j$ processes. At the LHC, $Wbj + X$ is dominant, while the other two backgrounds are roughly comparable.

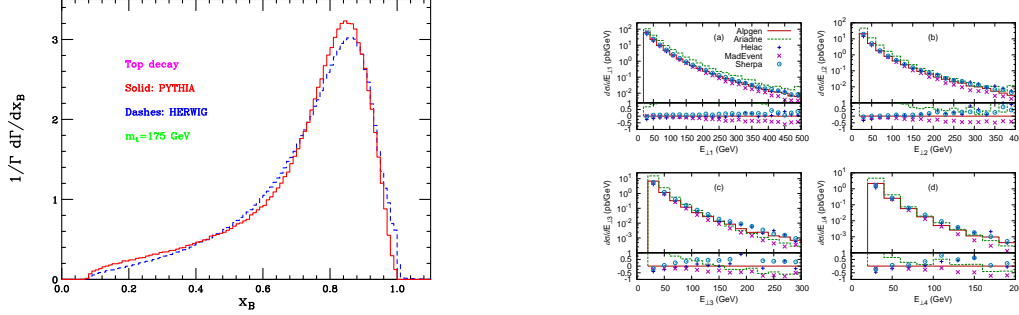


Fig. 5.: Left: B -hadron spectrum in top decay yielded by HERWIG and PYTHIA, after fitting their hadronization models to LEP and SLD data. Right: Transverse energy of the first four jets in $W + 4$ jet events according to a few generators.

We wish to present a few studies on the theoretical uncertainty on the top mass measurement. Ref. [33] discusses the error in the lepton+jets channel, due to different parametrizations of the underlying event in PYTHIA, with the possible inclusion of jet-energy-scale corrections. As showed in Fig. 4, several tunings and models are investigated and the overall uncertainty is estimated to be $\Delta m_t \simeq \pm 0.5$ GeV. Scale and parton-density uncertainties are instead studied in [34]. Fig. 4 also presents the NLO cross section $\sigma(pp \rightarrow tt)$, as a function of m_t , for various values of renormalization and factorization scales. The relation $\Delta m_t/m_t \sim 0.2 \Delta\sigma/\sigma + 0.05$ emphasizes that the relative error on m_t cannot be below 5%, if m_t is determined from a measurement of the cross section. Another source of uncertainty on m_t is b -quark fragmentation in top decay, which relies on the hadronization models implemented in MC generators, fitted to e^+e^- data. In Ref. [35] it was found that the default parametrizations of cluster and string models are incapable of fitting B -hadron data from LEP and SLD. After fitting few parameters, PYTHIA gives a good description of the data, whereas HERWIG is only marginally consistent. Therefore, as shown in Fig. 5, HERWIG and PYTHIA predictions for the B -spectrum in top decay look quite different: such a result will have an impact on the MC uncertainty on

TABLE I.: Cross sections in pb for $(Z/\gamma^* \rightarrow e^+e^-) + n$ jets at the LHC according to different codes. The following cuts are set on jets: $p_T > 20$ GeV, $|\eta| < 2.5$, $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} > 0.4$.

$Z + n$ jets	0	1	2	3	4	5	6
ALPGEN	723.4(9)	188.3(3)	69.9(3)	27.2(1)	10.95(5)	4.6(1)	1.85(1)
SHERPA	723.9(7)	189.6(9)	71.4(4)	30(2)			
CompHEP	730.9(1)	190.20(7)	70.22(7)				
MadEvent	723(1)	188.6(3)	69.3(1)	27.1(2)	10.6(1)		
GR@APPA	744(7)	182.77(8)	67.70(3)				

m_t . The analysis [36] investigates instead the systematic error due to initial- (ISR), final-state radiation (FSR), underlying event (UE) and hadronization on the reconstruction of m_t as the invariant mass of the Wb -jet combination, in the lepton+jets channel. The b -jet is clustered according to the k_T (KtJet) and cone (PxCone) algorithms. Ref. [36] found that, relying on the k_T algorithm, the reconstructed m_t shows visible dependence on FSR, ISR and UE, whereas hadronization effects are very little. When using the cone one, FSR and hadronization have a large impact, while ISR and UE effects are negligible. Given the different features of the two algorithms, using both of them is therefore advisable.

As for matrix-element generators, Table I [37] quotes the results of a few codes on $Z + n$ jets, with $n \leq 6$ and $Z \rightarrow e^+e^-$, one of the backgrounds for top production. It is remarkable that, as long as a process is implemented, the considered codes agree. More exclusive studies on $W +$ jets were carried out in [38], but even there the used programs were found to agree. Fig. 5 shows the transverse energy distribution of the first four hardest jets yielded by ALPGEN, ARIADNE, HELAC, MadEvent and SHERPA.

In summary, we reviewed some of the existing event generators for top signals and background and presented a few results for the LHC. Given the large numbers of available tools, it is always advisable using at least two codes for comparison.

REFERENCES

- [1] CORCELLA G. ET AL., *JHEP*, **0101** (2001) 010.
- [2] SJOSTRAND T., MRENNA S. and SKANDS P., *JHEP*, **0605** (2006) 026.
- [3] CORCELLA G. and SEYMOUR M.H., *Phys. Lett. B*, **442** (1998) 417.
- [4] CORCELLA G., MANGANO M.L. and SEYMOUR M.H., *JHEP*, **0007** (2000) 004.
- [5] NORRIN E. and SJOSTRAND T., *Nucl. Phys. B*, **603** (2001) 297.
- [6] CORCELLA G. and SEYMOUR M.H., *Nucl. Phys. B*, **565** (2000) 227.
- [7] SJOSTRAND T. and MIU G., *Phys. Lett. B*, **449** (1999) 313.
- [8] WEBBER B.R., *Nucl. Phys. B*, **238** (1984) 492.
- [9] ANDERSSON B. ET AL., *Phys. Rept.*, **97** (1983) 31.
- [10] SJOSTRAND T. and SKANDS P., *JHEP*, **0403** (2004) 053.
- [11] BUTTERWORTH J.M., FORSHAW J.R. and SEYMOUR M.H., *Z. Phys. C*, **72** (1996) 637.
- [12] BAHR M. ET AL., arXiv:0803.0883 [hep-ph].
- [13] SJOSTRAND T., MRENNA S. and SKANDS P., arXiv:0710.3820 [hep-ph].
- [14] LÖNNBLAD L., *Comput. Phys. Commun.*, **71** (1992) 15.
- [15] GLEISBERG T. ET AL., *JHEP*, **0402** (2004) 056.
- [16] CATANI S., KRAUSS F., KÜHN R. and WEBBER, B.R., *JHEP*, **0111** (2001) 063.
- [17] LÖNNBLAD L., *JHEP*, **0205** (2002) 046.
- [18] CATANI S. ET AL., *Nucl. Phys. B*, **406** (1993) 187.
- [19] FRIXIONE S. and WEBBER B.R., *JHEP*, **0206** (2002) 029.
- [20] FRIXIONE S., NASON P. and OLEARI C., *JHEP*, **0711** (2007) 70.
- [21] MANGANO M.L. ET AL., *JHEP*, **0307** (2003) 001.
- [22] STELZER T. and MALTONI F., *JHEP*, **0302** (2003) 027.
- [23] BOOS E. ET AL., *Nucl. Instrum. Meth. A*, **534** (2004) 250.
- [24] CAFARELLA A., PAPADOPOULOS C.G. and WOREK M., arXiv:0710.2427 [hep-ph].
- [25] TSUNO S. ET AL., *Comput. Phys. Commun.*, **175** (2006) 665.
- [26] KILIAN W., OHL T. and REUTER J., arXiv:0708.4233 [hep-ph].
- [27] FRIXIONE S., NASON P. and WEBBER B.R., *JHEP*, **0308** (2003) 007.
- [28] FRIXIONE S., NASON P. and RIDOLFI G., *JHEP*, **0709** (2007) 126.
- [29] MANGANO M.L., MORETTI M., PICCININI F. and TRECCANI M., *JHEP*, **0701** (2007) 013.
- [30] FRIXIONE S., MOTYLINSKI P., LAENEN E. and WEBBER B.R., *JHEP*, **0603** (2006) 092.

- [31] CAMPBELL J.M. and ELLIS R.K., <http://mcfm.fnal.gov/>.
- [32] CAMPBELL J.M. ET AL., *Phys. Rev. D*, **75** (2007) 054015.
- [33] SKANDS P. and WICKE D., *Eur. Phys. J. C*, **52** (2007) 133.
- [34] FREDERIX R. and MALTONI F., arXiv:0712.2355 [hep-ph].
- [35] CORCELLA G. and DROLLINGER V., *Nucl. Phys. B*, **730** (2005) 82.
- [36] SEYMOUR M.H. and TEVLIN C., *JHEP*, **0611** (2006) 052.
- [37] PICCININI F., private communication.
- [38] ALWALL J. ET AL., *Eur. Phys. J. C*, **53** (2008) 473.